

# Learning Autonomous Vehicle Safety Concepts from Demonstrations

Karen Leung, Sushant Veer, Edward Schmerling, Marco Pavone



### On the importance of having an independent safety evaluation module

Is the autonomous car safe?

Autonomous car      Human driver

Within the AV stack, there typically is a prediction model that predicts the behaviors of other agents, and a planner that uses the prediction model to make informed decisions. However, the prediction model is not always accurate, and the planner may not respond fast enough to split-second threats.

Therefore, there needs to be an independent safety module that can intervene with a safe maneuver anytime upstream components “make a mistake”. But *when* should such a safety module intervene, and *how* should it do so?

### What exactly is a “Safety Concept”?

First, we define some terminology to describe what our goal is.

Two functions mapping from world state to:

- Safety measure
- Set of allowable safe controls

World state → Safety measure (Unsafe/Safe) → Allowable safe controls

Examples:

- Velocity obstacles
- Safety Force Field
- Responsibility Sensitive Safety
- Forward reachability
- Backward reachability

### How can we synthesize a novel safety concept?

We can describe a family of safety concepts via HJ reachability

Hamilton-Jacobi-Isaacs partial differential equation (Robust HJB equation)

$$\frac{\partial V(x, t)}{\partial t} + \min \left\{ 0, \max_{u \in U} \min_{d \in D} \frac{\partial V(x, t)}{\partial x} \cdot f(x, u, d) \right\} = 0$$

$$V(x, 0) = v(x)$$

#### Hamilton-Jacobi Reachability

Open-loop “non-reactive” policies

Consider *all* possible behaviors  
Full forward reachable set

Consider only a *subset* of possible behaviors  
e.g., hard-braking (SFF)

Closed-loop “reactive” policies

Guard against *all* possible policies  
Including worst-case outcomes

Guard against a *subset* of possible policies  
Assumptions on other agent’s behaviors

By varying the parameters of the HJ reachability problem, we can describe both closed-loop and open-loop behaviors, and anything in between.

### How should we select “reasonable” control bounds?

Picking *all possible* controls & disturbances leads to overly conservative safety concepts

Given a dataset of states and controls:  $(x^{(1)}, u^{(1)}), (x^{(2)}, u^{(2)}), \dots, (x^{(N)}, u^{(N)})$  we want to learn  $U(x)$

**Key insight:** Humans take controls that keep them safe. Taking controls outside the boundary will lead to an undesirable outcome.

↓

Data lives inside a control invariant set.

### Control set learning via Control Barrier Functions

$$\max_{u \in U} \nabla b(x)^T f(x, u) \geq -\alpha(b(x)) \quad \forall x \in X$$

$b(x)$       Learn parameters of  $\alpha$  so that this condition holds for the dataset

Bound the rate at which the system approaches the boundary

### How do we account for the constraint coupling when synthesizing a safety concept with HJ reachability?

Safe interaction data → CBF learning  $U(x), D(x)$  → HJ reachability → Safety concept

**Proposition 1.** Consider a coupled affine constraint  $p(u_A, u_B) := au_A + bu_B + c \geq 0$  in  $u_A$  and  $u_B$ , and a linear objective  $q(u_A, u_B) = Au_A + Bu_B$ . Let  $\tilde{U}^A = \{u_A \in U^A \mid \exists u_B \in U^B, p(u_A, u_B) \geq 0\}$  and  $\tilde{U}^B = \{u_B \in U^B \mid \exists u_A \in U^A, p(u_A, u_B) \geq 0\}$  represent control sets for each agent which ensures the other agent can satisfy the constraint  $p(u_A, u_B) \geq 0$ . Let  $U^A(u_B) = \{u_A \in U^A \mid p(u_A, u_B) \geq 0\}$ , and  $U^B(u_A) = \{u_B \in U^B \mid p(u_A, u_B) \geq 0\}$  describe an agent’s feasible control set with the other agent’s control fixed. Then,

$$\max_{u_A \in \tilde{U}^A} \min_{u_B \in U^B(u_A)} q(u_A, u_B) \geq \min_{u_B \in \tilde{U}^B} \max_{u_A \in U^A(u_B)} q(u_A, u_B) \quad (9)$$

That is, the player that acts first has the advantage assuming the second player is provided feasible options.

Constrained min-max game

### What does a data-driven safety concept look like and what does it mean?

**Worst case analysis:** can be over-conservative

**Data-informed:** generated by propagating learned control bounds through dynamics

**Fixed policy (braking):** assumes too much; over-optimistic

Agent A

Agent B